

Universidade Federal do Paraná
Setor de Ciências Exatas
Departamento de Estatística
Programa de Especialização em *Data Science* e *Big Data*

Guilherme Câmara Cavalcanti Silva

**Indicadores fundamentalistas para explicar os
valores das ações dos principais bancos do
Brasil**

**Curitiba
2020**

Guilherme Câmara Cavalcanti Silva

Indicadores fundamentalistas para explicar os valores das ações dos principais bancos do Brasil

Monografia apresentada ao Programa de Especialização em Data Science e Big Data da Universidade Federal do Paraná como requisito parcial para a obtenção do grau de especialista.

Orientador: Prof. Fernando de Pol Mayer

Curitiba
2020

Indicadores fundamentalistas para explicar os valores das ações dos principais bancos do Brasil

Guilherme Câmara Cavalcanti Silva¹
Prof. Fernando de Pol Mayer²

Resumo

Foi realizado um desenvolvimento com base em preceitos da escola fundamentalista, e com base em dados de uma série histórica de 10 anos (2010 a 2019), com o objetivo de analisar quais são os indicadores financeiros e de mercado que mais explicam a valorização, ou desvalorização, dos preços das ações dos quatro maiores bancos com capital aberto do país (Banco do Brasil, Bradesco, Itaú e Santander). Para tanto, elaborou-se um modelo estatístico, mais precisamente um Modelo Linear Generalizado (MLG), que pode ser utilizado por investidores para identificar os preços-alvo das ações destes, e de outros bancos de mesma natureza, e indicar momentos propícios de compra e venda, baseado em determinados indicadores amplamente utilizados para avaliar a situação financeira de instituições bancárias.

Palavras-chave: indicadores fundamentalistas, setor bancário, modelo linear generalizado, cotação, ações.

Abstract

A development based on fundamentalist school precepts, and based on data from a 10-year historical series (2010 to 2019), with the main objective of analyzing which are the financial and market indicators that most explain the rise, or drop, share prices of the country's four largest publicly traded banks (Banco do Brasil, Bradesco, Itaú and Santander). For this purpose, a statistical model was elaborated, more precisely a Generalized Linear Model (GLM), which can be used by investors to identify the target prices of the shares of these and other banks of the same nature, and indicate suitable moments of purchase and sale, based on certain indicators widely used to assess the financial situation of banking institutions.

Keywords: fundamentalist indicators, banking sector, generalized linear model, price, stocks.

1 Introdução

Nos últimos 10 anos, o mercado de ações no Brasil, que é concentrado em apenas uma bolsa de valores, a B3, tem

crescido substancialmente em termos de volume de negócios. Este fato é facilmente corroborado pelo aumento do Ibovespa, principal índice da B3, configurando-se em sua definição como “o principal indicador de desempenho das ações negociadas na B3 e reúne as empresas mais importantes do mercado de capitais brasileiro.” (B3, 2020) [1], dentro deste período. Por sua vez, tal índice engloba em sua composição, dentre outras empresas, as maiores instituições bancárias brasileiras com capital aberto: Banco do Brasil, Bradesco, Itaú Unibanco e Santander. Os papéis destes bancos, em 14 de setembro de 2020, representavam, juntos, 15,17% da composição do índice, o que demonstra a relevância dos mesmos frente ao mercado acionário brasileiro.

Apresentando-se como uma oportunidade de diversificação de portfólio e de aumento de rentabilidade, o mercado acionário, em contrapartida, apresenta um risco inerente de perda financeira, característico dos investimentos em renda variável. Diante disso, diferentes metodologias são utilizadas por investidores para suportar as tomadas de decisão de compra e venda de ações, a fim de mitigar este risco inerente de perda. Entre estas metodologias, uma das mais difundidas e utilizadas é a análise fundamentalista, que “vê o mercado de ações (e a valorização dos seus ativos) como o resultado da atividade econômica desenvolvida por uma empresa” (DEBASTIANI, 2008, p.9) [2], em oposição à análise técnica, que foca seus esforços sobre as próprias volatilidades das ações.

Neste sentido, o presente trabalho tem como objetivo verificar, com base em preceitos da escola fundamentalista, e com base em dados de uma série histórica de 10 anos (2010 a 2019), quais são os indicadores financeiros e de mercado que mais explicam a valorização, ou desvalorização, dos preços das ações dos quatro maiores bancos com capital aberto do país, já supracitados. Além disso, pretende-se apresentar um modelo estatístico, mais precisamente um Modelo Linear Generalizado (MLG), que possa ser utilizado por investidores para identificar os preços-alvo das ações destes, e de outros bancos de mesma natureza, e indicar momentos propícios de compra e venda, baseado em dez indicadores amplamente utilizados para avaliar a situação financeira de instituições bancárias.

¹ Aluno do programa de Especialização em Data Science & Big Data, gccavalcantis@gmail.com.

² Professor do Departamento de Estatística - DEST/UFPR.

2 Contextualização e Motivações

Como citado na introdução, o mercado de ações, como uma possibilidade de investimento em renda variável, apresenta um risco inerente de perda financeira para os investidores. Por este motivo, é de suma importância que tais investidores utilizem técnicas e disponham de “ferramentas” que possibilitem a mitigação deste risco, e busquem maximizar os retornos que são esperados. Para os investidores que acreditam na escola fundamentalista, este trabalho apresenta-se com uma alternativa de utilização de um modelo estatístico, desenvolvido empiricamente, a ser utilizado para suportar as decisões de compra e venda de papéis de bancos comerciais.

Ademais, pretende-se apresentar, principalmente a investidores “pessoas físicas”, o racional de construção de modelos estatísticos para fins de investimento em ações, que normalmente são elaborados por empresas assessoras de investimentos.

Adicionalmente, uma outra motivação é a de gerar resultados, com base em análise descritiva e modelagem, que possam indicar quais são os indicadores que têm sido considerados como os mais relevantes para o mercado, quando na tomada de decisão para investir nestes bancos. Esta resposta mostra-se extremamente relevante aos investidores em bancos, frente a um novo cenário que tem se desenhado para o setor, de uma queda significativa de rentabilidade dos grandes bancos, ocasionados pela ascensão do open banking, de novas corretoras com modelo de negócio disruptivo e dos bancos digitais (fintechs).

Diante deste contexto, além da introdução e desta seção, este artigo contém os seguintes tópicos: metodologia e bases de dados, experimentos e análise de resultados, e conclusão.

3 Metodologia e bases de dados utilizadas

Anteriormente à etapa de obtenção de dados, foram observados alguns aspectos para delimitação da pesquisa realizada. Primeiramente, escolheu-se analisar apenas os papéis mais negociados na bolsa de valores de cada um dos bancos. Desta forma, foram os papéis analisados: BBAS3, BBAS3, ITUB4 e SANB11. Posteriormente, houve a delimitação da série histórica a ser analisada: dez anos, de 2010 a 2019. Este período em análise foi delimitado de acordo com a disponibilidade dos dados das demonstrações financeiras anuais de empresas com capital aberto, liberados publicamente pela CVM (Comissão de Valores Mobiliários) em seu portal público de dados, sendo que não foram disponibilizados dados anteriores à 2010 neste portal. Em seguida, delimitou-se que seriam consideradas apenas as cotações de fechamento dos papéis, nas datas de divulgação das demonstrações financeiras anuais, sempre no ano seguinte ao ano calen-

dário avaliado. Esta decisão foi tomada com base no fato de que são nestas datas de divulgação que o mercado reage mais fortemente, seja positivamente (comprando as ações), ou negativamente (vendendo as ações), com base nos indicadores que foram tornados públicos.

Quanto aos indicadores a serem considerados para análise, foram delimitados dez indicadores amplamente aceitos e utilizados por investidores adeptos à escola fundamentalista para avaliação da “saúde financeira” de bancos. São eles: patrimônio líquido, lucro líquido, receita de intermediação financeira, ROE (return on equity), margem líquida, PDD (provisão de devedores duvidosos), PDD/Lucro líquido, índice de basileia, proventos e payout.

Ademais, decidiu-se pela utilização do software R 4.0.2 [4] para execução da pesquisa e ajuste do modelo linear generalizado, por meio da função padrão `glm()`.

A partir destas definições, teve início a etapa de obtenção de dados. As cotações históricas foram obtidas em arquivos “.txt” disponíveis no site da B3 e importados no R utilizando o pacote `rb3` [5]. Já os dados necessários para cálculo dos indicadores financeiros e de mercado a serem analisados foram obtidos utilizando o pacote `GetCVMData` [3], que busca os dados de demonstrações financeiras no portal aberto de dados da CVM (Comissão de Valores Mobiliários). A partir da base completa de dados financeiros das empresas listadas na B3, foram filtrados os dados apenas dos bancos em escopo e calculados os indicadores.

A partir destas bases, foi montada uma base de dados, em que foram incluídos, nas colunas, os seguintes dados, para os dez anos em escopo: código da ação; nome do banco; cotações de fechamento dos papéis mais negociados de cada banco, nas datas de divulgação das demonstrações financeiras anuais, sempre no ano seguinte ao ano calendário avaliado e os indicadores; indicadores financeiros calculados com base nas demonstrações financeiras anuais dos bancos. A Tabela 1 mostra um resumo estatístico destas informações.

Uma vez estruturada a base de dados, deu-se início à etapa de análise descritiva e modelagem. Primeiramente, verificou-se as correlações existentes entre a cotação e os indicadores, por meio da geração e análise de matrizes de correlação e gráficos de dispersão. Desta forma, os indicadores foram classificados em ordem crescente, levando em conta os que apresentaram maior correlação com os preços das ações. Em seguida, dado o fato de que o principal objetivo do presente trabalho foi o de verificar quais indicadores têm maior impacto no comportamento do preço da ação, a cotação foi considerada a variável resposta e os indicadores as variáveis explicativas. Além disso, a variável “banco” foi considerada como uma das variáveis explicativas, por ser considerada uma variável de “controle”, já que naturalmente sabe-se que devem existir diferenças entre as cotações dos diferentes bancos.

Além da distribuição normal, foi testada também a distribuição gama para o modelo, uma vez que espera-se uma certa assimetria na variável resposta.

Tabela 1: Resumo estatístico dos principais indicadores analisados.

	Acao	Cotacao_Dia	PL	Lucro_Liquido	Margem_Liquida	ROE	Payout
X	BBDC4 :10	Min. :12.57	Min. : 41070763	Min. : 5459278	Min. :0.0300	Min. :0.06693	Min. :0.2138
X.1	SANB11:10	1st Qu.:25.18	1st Qu.: 75339348	1st Qu.: 9997601	1st Qu.:0.0875	1st Qu.:0.13155	1st Qu.:0.3250
X.2	ITUB4 :10	Median :33.34	Median : 85126298	Median :12768415	Median :0.1350	Median :0.16718	Median :0.3801
X.3	BBAS3 :10	Mean :31.46	Mean : 91313099	Mean :14255012	Mean :0.1237	Mean :0.15582	Mean :0.4223
X.4		3rd Qu.:38.09	3rd Qu.:104774780	3rd Qu.:17484134	3rd Qu.:0.1525	3rd Qu.:0.18625	3rd Qu.:0.4626
X.5		Max. :53.29	Max. :150466000	Max. :27813000	Max. :0.2000	Max. :0.22932	Max. :0.9399

Após os ajustes dos modelos utilizando os dois tipos diferentes de distribuição, houve aplicação do método “stepwise” para seleção do menor número de variáveis explicativas para a máxima explicação da variável resposta. Por fim foram feitas comparações entre os modelos otimizados, tendo sido analisados aspectos como: os valores de resíduos; comparativos dos modelos candidatos pelo AIC (*Akaike Information Criterion*) e pela razão de verossimilhanças entre eles. Após, definido o melhor modelo, foram realizadas previsões com base nos dados dos indicadores constantes na base de dados para prever os valores das cotações nas datas das cotações também contidas na base de dados. Foram analisados os resultados desta previsão e, adicionalmente, foram feitas outras duas previsões com base em valores hipotéticos extremos, contemplando dois cenários distintos, para, por fim, serem elaboradas as conclusões quanto à efetividade do modelo proposto.

4 Análise descritiva e modelagem

4.1 Análise de correlação entre cotação e indicadores

Para análise das correlações entre os valores das cotações e os valores dos indicadores financeiros e de mercado, primeiramente foi elaborada uma matriz de correlação, que não demonstrou nenhuma correlação de algum indicador com o preço da ação acima de 0,43 (Figura 1).

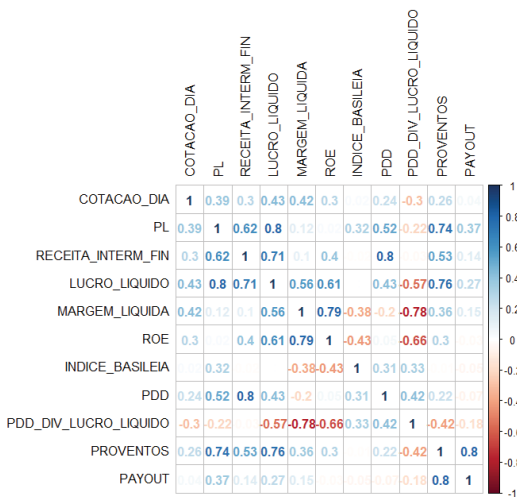


Figura 1: Matriz de correlação entre as variáveis analisadas

Além disso, foram gerados gráficos de dispersão (vide exemplo a partir da Figura 2) entre os valores de cotação (variável resposta) e os indicadores (variáveis explicativas), para cada um dos bancos. Na Figura 2, é mostrado um exemplo de gráfico, entre a cotação e o indicador “PL” (Patrimônio Líquido).

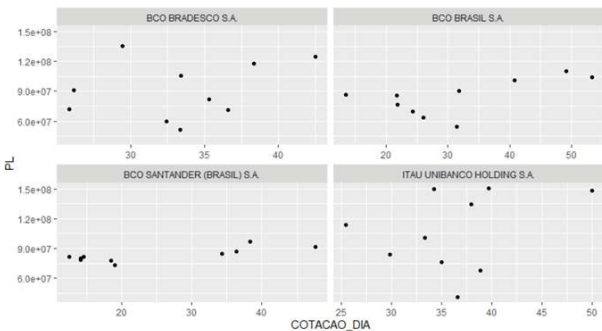


Figura 2: Exemplo de gráfico de dispersão entre a variável resposta e uma das variáveis explicativas

Após geração e análise das matrizes de correlação e dos gráficos de dispersão para todos os dez indicadores em análise, versus a cotação, definiu-se uma ordem de “relevância” dos indicadores com relação ao impacto no comportamento da ação, a ser aplicada no modelo, disposta da seguinte forma (do indicador de maior relevância, até o indicador de menor relevância): patrimônio líquido, lucro líquido, ROE, margem líquida, índice de basileia, PDD, PDD/lucro líquido, receita de intermediação financeira, payout e proventos.

4.2 Análise descritiva e modelagem

O primeiro passo da etapa de análise descritiva e modelagem foi gerar um histograma (Figura 3), com base nas cotações históricas da base dados, para definição sobre a distribuição a ser utilizada no modelo. Com base no gráfico gerado, decidiu-se por testar o modelo utilizando dois tipos de distribuição diferentes: normal e “gama”.

Em seguida, com base na ordem de relevância definida na etapa de análise de correlação entre cotação e indicadores, foram rodados os dois modelos candidatos, um considerando distribuição normal e outro gama. Logo após, aplicou-se o método “stepwise” para os dois modelos candidatos, com o objetivo de minimizar a quantidade de variáveis explicativas e maximizar a explicação da variável resposta.

Os resíduos de ambos os modelos mostraram-se satisfatórios, ou seja, ambos apresentaram uma dispersão

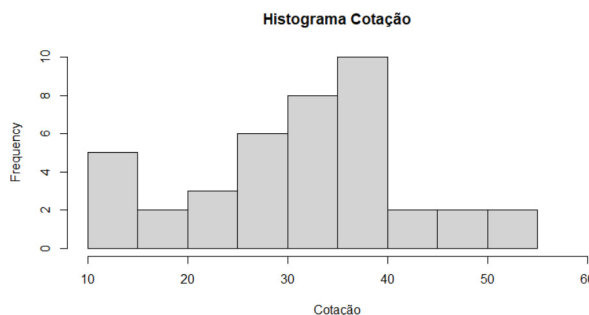


Figura 3: Histograma dos quarenta valores da variável resposta observados na base de dados

aleatória ao redor de zero e sem nenhuma tendência aparente (Figura 4).

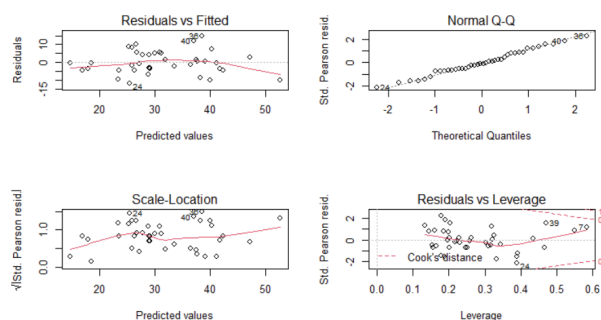


Figura 4: resultados observados para o modelo com distribuição normal

Após a aplicação do método "stepwise" para os dois modelos candidatos, verificou-se através do método "anova" que para o modelo de distribuição normal, apenas os indicadores patrimônio líquido, ROE e margem líquida revelaram-se significativos, enquanto para o modelo de distribuição gama, apenas os indicadores patrimônio líquido e ROE.

Entretanto, ao comparar-se os modelos candidatos pelo AIC, o modelo com distribuição normal apresentou o menor valor, indicando assim ser mais adequado. Desta forma, optou-se pelo modelo utilizando distribuição normal para continuidade e execução de testes e ajustes (Tabela 2).

Tabela 2: Comparação AIC modelos.

	df	AIC
MOD_NORMAL	12.00	283.56
MOD_GAMA	7.00	292.09

Feita a opção pelo modelo com distribuição normal (após aplicação do método "stepwise"), foi feita a análise do método "anova" e observados os seguintes resultados, apresentados na Tabela 3.

Com base nestes resultados, foram feitos testes do modelo, primeiramente retirando o indicador lucro líquido, e posteriormente retirando o indicador PDD. Ao retirar o lucro líquido, percebeu-se que esta, apesar de não ter sido indicada como uma variável significativa de início,

precisa estar no modelo, o que foi concluído após análise de AIC entre o modelo sem o indicador lucro líquido e o modelo anterior com o indicador lucro líquido.

Além disso, ao aplicar o método "anova" entre os dois modelos, percebeu-se uma deviance significativa, reforçando que o indicador lucro líquido deveria permanecer no modelo (Tabela 4).

Já ao retirar o indicador PDD do modelo após o "stepwise", o AIC mostrou-se menor e a redução na deviance na comparação entre os dois modelos (com e sem PDD) não se mostrou significativa. Aplicando o método "anova" no modelo sem PDD, o indicador PDD/Lucro líquido mostrou-se não significativo, com p valor = 0.826. Desta forma, testou-se o modelo novamente, sem este indicador. Por fim, chegou-se ao modelo definitivo (Tabela 5).

4.3 Predição

Após a conclusão da etapa de análise descritiva e modelagem, e definição do modelo a ser utilizado, foi realizada uma etapa de predição, com base nos valores dos indicadores da base de dados utilizada para aprimorar o modelo. Ao gerar o gráfico com os valores reais das cotações da base de dados e os valores preditos, observou-se o resultado demonstrado no gráfico exposto na Figura 5, que demonstra uma relação de linearidade entre os valores reais e os preditos, o que significa um resultado satisfatório na utilização do modelo para prever os valores das ações.

Além disso, com base neste modelo, foram simulados alguns cenários hipotéticos de flutuação dos indicadores, com o objetivo de verificar qual seria o comportamento do preço da ação. A fim de simular um possível cenário que pode vir a se desenhar no setor bancário nos próximos anos, de queda de rentabilidade dos grandes bancos, dado o advento das "fintechs" e do "open banking", como já comentado na seção de motivações deste artigo, foi realizada uma predição com base no seguinte cenário (a partir dos últimos dados disponíveis, verificados para o ano de 2019): aumento de 4% de patrimônio líquido, aumento de apenas 2% do lucro líquido, queda de 2% do ROE, queda de 3% da margem líquida e aumento do percentual de payout pago, com a adição de 5 pontos percentuais.

Tabela 3: Anova do modelo de distribuição normal

	Df	Deviance	Resid. Df	Resid. Dev	F	Pr(>F)
NULL			39	4248.31		
ACAO	3	672.67	36	3575.65	4.22	0.0136
PL	1	365.33	35	3210.32	6.88	0.0138
LUCRO_LIQUIDO	1	3.97	34	3206.36	0.07	0.7866
ROE	1	1031.96	33	2174.40	19.43	0.0001
MARGEM_LIQUIDA	1	235.79	32	1938.60	4.44	0.0439
PDD	1	53.64	31	1884.97	1.01	0.3233
PDD_DIV_LUCRO_LIQUIDO	1	191.10	30	1693.86	3.60	0.0679
PAYOUT	1	153.27	29	1540.59	2.89	0.1001

Tabela 4: Anova dos modelos com e sem lucro líquido

	Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
1	30	2447.97			
2	29	1540.59	1	907.38	0.0000

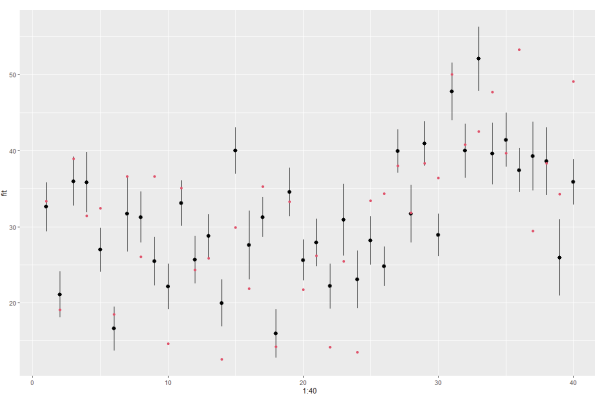


Figura 5: Comparação dos valores reais e preditos (com respectivos erros padrões)

Se comparados com os valores das cotações do ano de 2019, as cotações dos bancos apresentaram as seguintes variações: para o Bradesco, valorização de 39,5%; para o Santander, estabilização do preço (desvalorização de menos de 0,1%); para o Itaú, desvalorização de 20,6%; para o Banco do Brasil, desvalorização de 26,4%. Isso demonstra que para cada um dos bancos, a relevância dos indicadores financeiros possuem impacto totalmente diferente para determinação do preço da ação.

Posteriormente, foi realizada uma nova predição com base em um outro cenário hipotético, ainda pessimista, mas com os valores inversos do cenário hipotético anterior: diminuição de 4% de patrimônio líquido, diminuição de 2% do lucro líquido, aumento de 2% do ROE, aumento de 3% da margem líquida e diminuição do percentual de payout pago, com a subtração de 5 pontos percentuais.

Se comparados com os valores das cotações do ano de 2019, as cotações dos bancos apresentaram as seguintes variações: para o Bradesco, valorização de 91%; para o Santander, valorização de 40,4%; para o Itaú, valorização de 26,1%; para o Banco do Brasil, valorização de 2%. Isso demonstra que para todos os bancos, mais importantes que os aumentos de patrimônio líquido e lucro líquido, são os valores de margem líquida e ROE, o que faz sentido, visto que esses indicadores geralmente

são os mais levados em conta pelos investidores, pois retratam a performance de gestão dos bancos.

5 Conclusões

Os resultados obtidos indicam, primeiramente, que, dos dez indicadores financeiros e de mercado que são geralmente analisados por investidores para justificar os movimentos de compra e venda de papéis de bancos, apenas cinco deles remaneceram no modelo definitivo, por terem se mostrado significativos para explicarem as cotações das ações observadas. Isso pode indicar que, por mais que alguns destes indicadores sejam recomendados para análise por algumas empresas assessoras de investimentos, alguns deles podem distorcer uma decisão de compra ou venda destes papeis.

Ademais, após predição do modelo com base em cenários hipotéticos, fica claro que os indicadores que mais explicaram os preços das ações, para os quatro bancos analisados, são a margem líquida e o ROE, que, de fato, são os principais indicadores de performance de gestão financeira de bancos que são considerados como relevantes pelo mercado. Este resultado poderia servir de diretriz aos bancos para busca de maiores margens, ao invés de maiores receitas e/ou aumento de portfólio de produtos. Adicionalmente, o ROE como indicador significativo revela que os bancos só devem realizar investimentos que realmente sejam muito facilmente percebidos pelo mercado como necessários para o aumento substancial do lucro líquido.

Em suma, de acordo com o modelo apresentado, o que o mercado tem, historicamente, priorizado na hora de investir nos bancos é o fator eficiência - o que pode indicar uma necessidade de redução de custos e foco em operações lucrativas nos próximos anos. Em contrapartida, frente ao cenário do advento das “fintechs” e do “open banking”, os grandes bancos precisarão seguir um caminho inevitável de inovação, para não ficarem ultrapassados e serem substituídos por empresas que ofereçam novas soluções disruptivas, mais atraentes aos clientes. Desta forma, os grandes bancos precisarão implementar estes novos negócios com muita eficiência, de

Tabela 5: Anova modelo definitivo

	Df	Deviance	Resid. Df	Resid. Dev	F	Pr(>F)
NULL			39	4248.31		
ACAO	3	672.67	36	3575.65	4.17	0.0137
PL	1	365.33	35	3210.32	6.79	0.0140
LUCRO_LIQUIDO	1	3.97	34	3206.36	0.07	0.7878
ROE	1	1031.96	33	2174.40	19.18	0.0001
MARGEM_LIQUIDA	1	235.79	32	1938.60	4.38	0.0446
PAYOUT	1	270.68	31	1667.92	5.03	0.0322

forma que a margem e o ROE não sejam muito impactados. Como os “spreads” dessas operações disruptivas tendem a ser menor do que os obtidos através dos serviços tradicionalmente prestados pelos grandes bancos, há um indicativo de que os próximos anos serão desafiadores para os bancos, o que pode também significar uma estabilização, ao menos temporária, dos valores de mercado. Por outro lado, isto não pode ser apontado como uma certeza, visto que dados os seus tamanhos, rentabilidades e estruturas, os grandes bancos possuem alto poder de “dominação” do mercado deste setor. Por exemplo, o Itaú comprou, nos últimos anos, uma fatia de 46% da XP investimentos, corretora de investimentos que tem crescido exponencialmente, o que ilustra muito bem a capacidade dos grandes bancos de continuarem dominando o mercado.

- [6] T. Reis, *Os indicadores mais importantes em uma análise*, (Sunoresearch, 2017. Disponível em: <https://www.sunoresearch.com.br/artigos/os-indicadores-mais-importantes-em-uma-analise>. Acesso em: 05 de jul. de 2020).

Agradecimentos

Gostaria de agradecer aos professores dos departamentos de Estatística, Matemática e Informática da Universidade Federal do Paraná, em especial ao professor Fernando de Pol Mayer, meu orientador no desenvolvimento deste artigo.

Referências

- [1] B3. *Índice Bovespa (Ibovespa)*. B3, São Paulo, Brasil. URL http://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm, 2020.
- [2] C. A. Debastiani and F. A. Russo, *Avaliando Empresas, Investindo em Ações: A aplicação prática da análise fundamentalista na avaliação de empresas*, (Novatec Editora, 2008), 1ª Edição.
- [3] Perlin, M and Kirch, G. *GetCVMDData: Reading Annual and Quarterly Financial Reports from CVM*. R package version 0.5. <https://github.com/msperlin/GetCVMDData>, 2020.
- [4] R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 2020.
- [5] Freitas, W. *rb3: Read public files released by B3 Exchange*. R package version 0.0.1.9000. <https://github.com/wilsonfreitas/rb3>, 2020.